AD NO. ___________________________
DTC PROJECT NO. 8-CO-160-UXO-021
REPORT NO. ATC 10336

STANDARDIZED

UXO TECHNOLOGY DEMONSTRATION SITE

SCORING RECORD NO. 924

SITE LOCATION:
ABERDEEN PROVING GROUND

DEMONSTRATOR:
SCIENCE APPLICATIONS INTERNATIONAL
CORPORATION (SAIC)
104 CLEMATIS AVENUE
WALTHAM, MA 02453

TECHNOLOGY TYPE/PLATFORM:
SIMPLIFIED COMBINED EMI MAGNETOMETER
PROTOTYPE (SCEMP)
DUAL MODE/PUSHCART

AREA COVERED:
INDIRECT FIRE

PREPARED BY:
U.S. ARMY ABERDEEN TEST CENTER
ABERDEEN PROVING GROUND, MD 21005-5059

OCTOBER 2010

Prepared for:
SERDP/ESTCP
MUNITIONS MANAGEMENT
ARLINGTON, VA 22203

U.S. ARMY DEVELOPMENTAL TEST COMMAND
ABERDEEN PROVING GROUND, MD 21005-5055

DISTRIBUTION UNLIMITED, OCTOBER 2010.

DISPOSITION INSTRUCTIONS

REPORT DOCUMENTATION PAGE						Form Approved OMB No. 0704-0188

1. REPORT DATE *(DD-MM-YYYY)*
August 2010

2. REPORT TYPE
Final

3. DATES COVERED *(From - To)*
16 through 18 November 2009

4. TITLE AND SUBTITLE
STANDARDIZED UXO TECHNOLOGY DEMONSTRATION SITE SCORING RECORD NO. 924

5a. CONTRACT NUMBER

5b. GRANT NUMBER

5c. PROGRAM ELEMENT NUMBER

6. AUTHOR(S)
McClung, Stephen J.

5d. PROJECT NUMBER
8-CO-160-UXO-021

5e. TASK NUMBER

5f. WORK UNIT NUMBER

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)
U.S. Army Developmental Test Command
400 Colleran Road
Aberdeen Proving Ground, MD 21005-5059

8. PERFORMING ORGANIZATION REPORT NUMBER
ATC-10336

9. SPONSORING/MONITORING AGENCY NAME(S) AND ADDRESS(ES)
Program Manager
SERDP/ESTCP
Munitions Management
Arlington, VA 22203

10. SPONSOR/MONITOR'S ACRONYM(S)

11. SPONSOR/MONITOR'S REPORT NUMBER(S)
Same as Item 8

12. DISTRIBUTION/AVAILABILITY STATEMENT
Distribution unlimited.

13. SUPPLEMENTARY NOTES
None

14. ABSTRACT
This scoring record documents the efforts of the Environmental Security Technology Certification Program to detect and discriminate inert unexploded ordnance (UXO) utilizing the APG Standardized UXO Technology Demonstration Site calibration lanes and open field sites. This Scoring Record was coordinated by J. Stephen McClung and the Standardized UXO Technology Demonstration Site Scoring Committee. Organizations on the committee include the U.S. Army Corps of Engineers, the Environmental Security Technology Certification Program, the Strategic Environmental Research and Development Program, the Institute for Defense Analysis, the U.S. Army Environmental Command, and the U.S. Army Aberdeen Test Center.

15. SUBJECT TERMS
Environmental Security Technology Certification Program, APG, Standardized UXO Technology Demonstration Site, open field, indirect fire, dual mode, pushcart.

16. SECURITY CLASSIFICATION OF:			17. LIMITATION OF ABSTRACT	18. NUMBER OF PAGES	19a. NAME OF RESPONSIBLE PERSON
a. REPORT	b. ABSTRACT	c. THIS PAGE			19b. TELEPHONE NUMBER *(Include area code)*
Unclassified	Unclassified	Unclassified	SAR		

ACKNOWLEDGMENTS

Authors:

Leonardo Lombardo
J. Stephen McClung Jr.
William Burch
Homeland Defense and Sustainment Division (HDSD)
U.S. Army Aberdeen Test Center (ATC)
Aberdeen Proving Ground (APG)

Rick Fling
Aberdeen Test Support Services (ATSS)
Sverdrup Technology, Inc.
Aberdeen Proving Ground

Christina McClung
U.S. Army Aberdeen Test Center
Survivability/Lethality Directorate
Aberdeen Proving Ground

TABLE OF CONTENTS

SECTION 4. TECHNICAL PERFORMANCE RESULTS

SECTION 5. APPENDIXES

SECTION 1. GENERAL INFORMATION

1.1 BACKGROUND

Technologies under development for the detection and discrimination of military munitions (MM) (i.e. unexploded ordnance {UXO} and discarded military munitions {DMM}) require testing so that performance can be characterized. To that end, Standardized Test Sites have been developed at Aberdeen Proving Ground (APG), Maryland, and U.S. Army Yuma Proving Ground (YPG), Arizona. These test sites provide a diversity of geology, climate, terrain, and weather as well as diversity in munitions and clutter. Testing at these sites is independently administered and analyzed by the government for the purposes of characterizing technologies, tracking performance with system development, comparing performance of different systems, and comparing performance in different environments (ref 1).

The Standardized UXO Technology Demonstration Site Program is a multiagency program spearheaded by the U.S. Army Environmental Command (USAEC). The U.S. Army Aberdeen Test Center (ATC) and the U.S. Army Corps of Engineers Engineering Research and Development Center (ERDC) provide programmatic support. The program is being funded and supported by the Environmental Security Technology Certification Program (ESTCP), the Strategic Environmental Research and Development Program (SERDP), and the U.S. Army Environmental Quality Technology (EQT) Program.

1.2 SCORING OBJECTIVES

The objective in the Standardized UXO Technology Demonstration Site Program is to evaluate the detection and discrimination capabilities of a given technology under various field and soil conditions. Inert munitions and clutter items are positioned in various orientations and depths in the ground.

The evaluation objectives are as follows:

a. To determine detection and discrimination effectiveness under realistic scenarios with various targets, geology, clutter, density, topography, and vegetation.

b. To determine cost, time, and workforce requirements to operate the technology.

c. To determine the demonstrator's ability to analyze survey data in a timely manner and provide prioritized Target Lists with associated confidence levels.

d. To provide independent site management to enable the collection of high quality, ground-truth (GT), geo-referenced data for post-demonstration analysis.

1.2.1 <u>Scoring Methodology</u>

a. The scoring of the demonstrator's performance is conducted in two stages: response stage and discrimination stage. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver-operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of clutter detection (P_{cd}) or the probability of false positive (P_{fp}). Those that do not correspond to any known item are termed background alarms. The background alarms are addressed as either probability of background alarm (P_{ba}) or background alarm rate (BAR).

b. The response stage scoring evaluates the ability of the system to detect emplaced targets without regard to ability to discriminate munitions from other anomaly sources. For the blind grid response stage, the demonstrator provides a target response from each and every grid square along with a threshold below which target responses are deemed insufficient to warrant further investigation. This list is generated with minimal processing and, since a value is provided for every grid square, includes amplitudes both above and below the system noise level. For the open field, the demonstrator provides a list of all anomalies deemed to exceed a demonstrator selected target detection threshold. An item (either munition or clutter) is counted as detected if a demonstrator indicates an anomaly within a specified distance (Halo Radius (R_{halo})) of a ground truth item.

c. The discrimination stage evaluates the demonstrator's ability to correctly identify munitions as such and to reject clutter. For the blind grid discrimination stage, the demonstrator provides the output of the discrimination stage processing for each grid square. For the open field, the demonstrator provides the output of the discrimination stage processing for anomaly reported in the response stage. The values in these lists are prioritized based on the demonstrator's determination that a location is likely to contain munitions. Thus, higher output values are indicative of higher confidence that a munitions item is present at the specified location. For digital signal processing, priority ranking is based on algorithm output. For other discrimination approaches, priority ranking may be based on rule sets or human judgment. The demonstrator also specifies the threshold in the prioritized ranking that provides optimum performance, (i.e., that is expected to retain all detected munitions and reject the maximum amount of clutter).

d. The demonstrator is also scored on efficiency and rejection ratios, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of munitions detections from the anomaly list, while rejecting the maximum number of anomalies arising from nonmunitions items. Efficiency measures the fraction of detected munitions retained after discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the maximum number of munitions detectable by the sensor and its accompanying clutter detection/false positive rate or BAR.

e. Based on configuration of the GT at the standardized sites and the defined scoring methodology, in some cases, there exists the possibility of having anomalies within overlapping halos and/or multiple anomalies within halos. In these cases, the following scoring logic is implemented:

(1) In situations where multiple anomalies exist within a single R_{halo}, the anomaly with the strongest response or highest ranking will be assigned to that particular GT item. If the responses or rankings are equal, then the anomaly closest to the GT item will be assigned to the GT item. Remaining anomalies are retained and scored until all matching is complete.

(2) Anomalies located within any R_{halo} that do not get associated with a particular GT item are excess alarms and will be disregarded.

f. In some cases, groups of closely spaced munitions have overlapping halos. The following scoring logic is implemented (App A, fig. A-1 through A-9):

(1) Overall site scores (i.e., P_d) will consider only isolated munitions and clutter items.

(2) GT items that have overlapping halos (both munitions and clutter) will form a group and groups may form chains.

(3) Groups will have a complex halos composed of the composite halos of all its GT items.

(4) Groups will have three scoring factors: groups found, groups identified, and group coverage. Scores will be based on 1:1 matches of anomalies and GT.

(a) Groups Found (Found): the number of groups that have one or more GT items matched divided by the total number of groups. Demonstrators will be credited with detecting a group if any item within the group is matched to an anomaly in their lists.

(b) Groups Identified (ID): the number of groups that have two or more GT items matched divided by the total number of groups. Demonstrators will be credited with identifying that a group is present if multiple items within the composite halo are matched to anomalies in their lists.

(c) Group Coverage (Coverage): the number of GT items matched within groups divided by the total number of GT items within groups. This metric measures the demonstrator accuracy in determining the number of anomalies within a group. If five items are present and only two anomalies are matched, the demonstrator will score 0.4. If all five are matched, the demonstrator will score 1.0.

(5) Location error will not be reported for groups.

(6) Demonstrators will not be asked to call out groups in their scoring submissions. If multiple anomalies are indicated in a small area, the demonstrator will report all individual anomalies.

(7) Excess alarms within a halo will be disregarded.

g. All scoring factors are generated utilizing the Standardized UXO Probability and Plot Program, version 4.

1.2.2 Scoring Factors

Factors to be measured and evaluated as part of this demonstration include:

a. Response stage ROC curves:

(1) Probability of detection (P_d^{res}).

(2) Probability of clutter detection (P_{cd}).

(3) Background alarm rate (BAR^{res}) or probability of background alarm (P_{ba}^{res}).

b. Discrimination stage ROC curves:

(1) Probability of detection (P_d^{disc}).

(2) Probability of false positive (P_{fp}).

(3) Background alarm rate (BAR^{disc}) or probability of background alarm (P_{ba}^{disc}).

c. Metrics:

(1) Efficiency (E).

(2) False positive rejection rate (R_{fp}).

(3) Background alarm rejection rate (R_{ba}).

d. Other:

(1) Probability of detection by size, depth, and density.

(2) Classification by type (i.e., 20-, 40-, 105-mm, etc.).

(3) Location accuracy for single munitions.

(4) Equipment setup, calibration time, and corresponding worker-hour requirements.

(5) Survey time and corresponding worker-hour requirements.

(6) Reacquisition/resurvey time and worker-hour requirements (if any).

(7) Downtime due to system malfunctions and maintenance requirements.

SECTION 2.   DEMONSTRATION

2.1   DEMONSTRATOR INFORMATION

2.1.1   Demonstrator Point of Contact (POC) and Address

POC: Mr. Robert Siegel
617-618-4662

Address: Science Applications International Corporation (SAIC)
104 Clematis Avenue
Waltham, MA   02453

2.1.2   System Description (provided by demonstrator)

The Simplified Combined EMI Magnetometer Prototype (SCEMP) as shown in Figure 1, is a man-pushed instrument utilizing both a total field magnetometer and an EM61 pulsed induction sensor. One Geonics EM61 Mk2A pulsed induction sensor, with one 1 x 1/2-meter coil 15 inches above ground, is employed. One Geometrics 822A cesium vapor magnetometer is deployed 4 feet from the edge of the EM61 coil, at a height of 18 inches above ground. A GPS receiver is employed to position the two sensors. Custom hardware interleaves the magnetometer data between the EM61 pulses. Forward rate of advance is approximately 1 meter per second.

Figure 1.   SAIC SCEMP/pushcart.

Support equipment required: A Global Positioning System (GPS) base station is used, and is set up over a known geodetic coordinate. Although SCEMP can be quickly taken apart, we would like to be able to store it intact overnight inside the building at the test site. We also need to charge batteries overnight inside the building.

Frequency and radio utilization: A Trimble TrimMark 3 GPS radio is used by the GPS base station to communicate with the GPS rover on the vehicle. Thirteen frequencies are available in the 461- to 464-MHz UHF band.

2.1.3 Data Processing Description (provided by demonstrator)

Target selection criteria: This section will detail the target selection criteria and the data required to implement the criteria by answering the following questions:

a. What kind of pre-processing (if any) is applied to the raw data (e.g., filtering, etc.)? A notch filter is applied to the magnetometer data to remove 60-Hz hum. A de-median filter is applied to both the magnetometer and EM61 data to remove diurnal drift, instrument drift, and the background from the platform itself.

b. What is the format of the data both pre- and post-processing of the raw data (e.g., ASCII, binary, etc.)? Raw data are stored in a binary format. Pre-processed data are stored in a Geosoft Oasis Montaj comma-delimited (.csv) format.

c. What algorithm is used for detection (e.g., peaks of signal surpassing threshold, etc.)? Peak detection.

d. Why is this algorithm used and not others? Simplicity.

e. On what principles is the algorithm based (e.g., statistical models, heuristic rules, etc.)? NA.

f. What tunable parameters (if any) are used in the detection process (e.g., threshold on signal amplitude, window length, filter coefficients, etc)? NA.

g. What are the final values of all tunable parameters for the detection algorithm? Depends on the noise levels at the site and the targets to be detected.

Parameter estimation: This section should include the details of which parameters will be extracted from the sensor data for each detected item for characterization. Please answer the following questions:

a. Which characteristics will be extracted from each detected item and input to the discrimination algorithm (e.g., depth, size, polarizability coefficients, fit quality, etc.)? NA.

b. Why have these characteristics been chosen and not others (e.g., empirical evidence of their ability to help discriminate, inclusion in a theoretical tradition, etc.)? NA.

c. How are these characteristics estimated (e.g., least-mean-squares fit to a dipole model, etc.; include the equations that are used for parameter estimation)? NA.

d. What tunable parameters (if any) are used in the characterization process? (e.g., thresholds on background noise, etc.)? NA.

Classification: This section should include the details describing the algorithm and associated data and parameters used for discrimination by answering the following questions:

a. What algorithm is used for discrimination (e.g., multi-layer perception, support vector machine, etc.)? NA.

b. Why is this algorithm used and not others? NA.

c. Which parameters are considered as possible inputs to the algorithm? NA.

d. What are the outputs of the algorithm (probabilities, confidence levels)? NA.

e. How is the threshold set to decide where the munitions/non-munitions line lies in the discrimination process? NA.

Training: This section should include the details of how training data are used to make a decision on the likelihood of the anomaly correspondence to munitions. Please answer the following questions: NA.

a. Which tunable parameters have final values that are optimized over a training set of data and which have values that are set according to geophysical knowledge (i.e., intuition, experience, common sense)? NA.

(1) For those tunable parameters with final values set according to geophysical knowledge:

(a) What is the reasoning behind choosing these particular values? NA.

(b) Why were the final values not optimized over a training set of data? NA.

(2) For those tunable parameters with final values optimized over the training set data:

(a) What training data are used (e.g., all data, a randomly chosen portion of data, etc.)? NA.

(b) What error metric is minimized during training (e.g., mean squared error, etc.)? NA.

(c) What learning rule is used during training (e.g., gradient descent, etc.)? NA.

(d) What criterion is used to stop training (e.g., number of iterations exceeds threshold, good generalization over validation set of data, etc.)? NA.

(e) Are all tunable parameters optimized at once or in sequence (“in sequence” = parameter 1 is held constant at some common sense value while parameter 2 is optimized, and then parameter 2 is held constant at its optimized value while parameter 1 is optimized)? NA.

b. What are the final values of all tunable parameters for the characterization process? NA.

2.1.4 Data Submission Format

Data were submitted for scoring in accordance with data submission protocols outlined on the USAEC Web site www.uxotestsites.org. These submitted data are not included in this report in order to protect GT information.

2.1.5 Demonstrator Quality Assurance (QA) and Quality Control (QC) (provided by demonstrator)

Overview of QC: This section is an overview of the complete QC portion of the QA/QC plan. The QC portion is the description of how systems checks are done by the demonstrator to check on items such as tracking, accuracy, drift, and system performance.

The GPS base station is set up using a checklist. Another known geodetic location is then occupied to verify that its coordinates are correct. The system is turned on and warmed up for 20 minutes. A ferrous object is put in front of each sensor in the array to ensure that the sensors are mapped to the correct channels. A 5-minute static test is conducted at the beginning and end of each day to ensure that noise levels are nominal. Instrument drift is corrected via background leveling in post-processing (generally a demedian filter).

Overview of QA: This section is an overview of the complete QA portion of the QA/QC plan. The QA portion is the description of the procedures to be employed during the demonstration to include items such as lane spacing, sampling rates, and estimated accuracy of navigation and tracking systems.

Survey lines are run by aligning one of the tires with the tire track on the previous line; this ensures that the inboard sensor is overlapped with the outboard sensor from the prior track. If tire tracks are not visible due to rocky terrain, a Trimble agricultural-grade track guidance system (parallel swathing) is employed.

2.1.6 Additional Records

The following record(s) by this vendor can be accessed via the Internet as Microsoft Word documents at www.uxotestsites.org.

2.2 APG SITE INFORMATION

2.2.1 Location

The APG Standardized Test Site is located within a secured range area of the Aberdeen Area. The Aberdeen Area of APG is located approximately 30 miles northeast of Baltimore at the northern end of the Chesapeake Bay. The Standardized Test Site encompasses 17 acres of upland and lowland flats, woods, and wetlands.

2.2.2 Soil Type

According to the soils survey conducted for the entire area of APG in 1998, the test site consists primarily of Elkton Series type soil (ref 2). The Elkton Series consist of very deep, slowly permeable, poorly drained soils. These soils formed in silty aeolin sediments and the underlying loamy alluvial and marine sediments. They are on upland and lowland flats and in depressions of the Mid-Atlantic Coastal Plain. Slopes range from 0 to 2 percent.

ERDC conducted a site-specific analysis in May 2002 (ref 3). The results basically matched the soil survey mentioned above. Seventy percent of the samples taken were classified as silty loam. The majority (77 percent) of the soil samples had a measured water content between 15 and 30 percent with the water content decreasing slightly with depth.

For more details concerning the soil properties at the APG test site, go to www.uxotestsites.org on the Web to view the entire soils description report.

2.2.3 Test Areas

A description of the test site areas at APG is presented in Table 1. A test site layout is shown in Figure 2.

TABLE 1. TEST SITE AREAS

Area	Description
Calibration lanes	Contains 14 standard munitions items buried in six positions, with representation of clutter, at various angles and depths to allow demonstrators to calibrate their equipment.
Blind grid	Contains 400 grid cells in a 0.5-acre site. The center of each grid cell contains either munitions, clutter, or nothing.
Open field	A 10-acre site composed of generally open and flat terrain with minimal clutter and minor navigational obstacles. Vegetation height varies from 15 to 25 cm. This area is subdivided into four subareas (legacy, direct fire, indirect fire, and challenge).
	• *Open field (legacy)* The legacy subarea contains the same wide variety of randomly-placed munitions that were present in the open field prior to the January 2008 general reconfiguration of the site.
	• *Open field (direct fire)* The direct fire subarea contains only three munition types that could be typically found at an impact area of a direct fire weapons range. Munitions and clutter are placed in a pattern typical for these munitions.
	• *Open field (indirect fire)* The indirect fire subarea contains only three munition types that could be typically found at an impact area of an indirect fire weapons range. Munitions and clutter are placed in a pattern typical for these munitions.
	• *Open field (challenge)* The challenge subarea is easily reconfigurable to meet the specific needs and requirements of the demonstrator or the program sponsor. Any results from this area are not reported in the standardized scoring record.
Woods	1.34-acre area consisting of cleared woods (tree removal with only stumps remaining), partially cleared woods (including all underbrush and fallen trees), and virgin woods (i.e., woods in natural state with all trees, underbrush, and fallen trees left in place).
Moguls	1.30-acre area consisting of two areas (the rectangular or driving portion of the course and the triangular section with more difficult, nondrivable terrain). A series of craters (as deep as 0.91 m) and mounds (as high as 0.91 m) encompass this section.

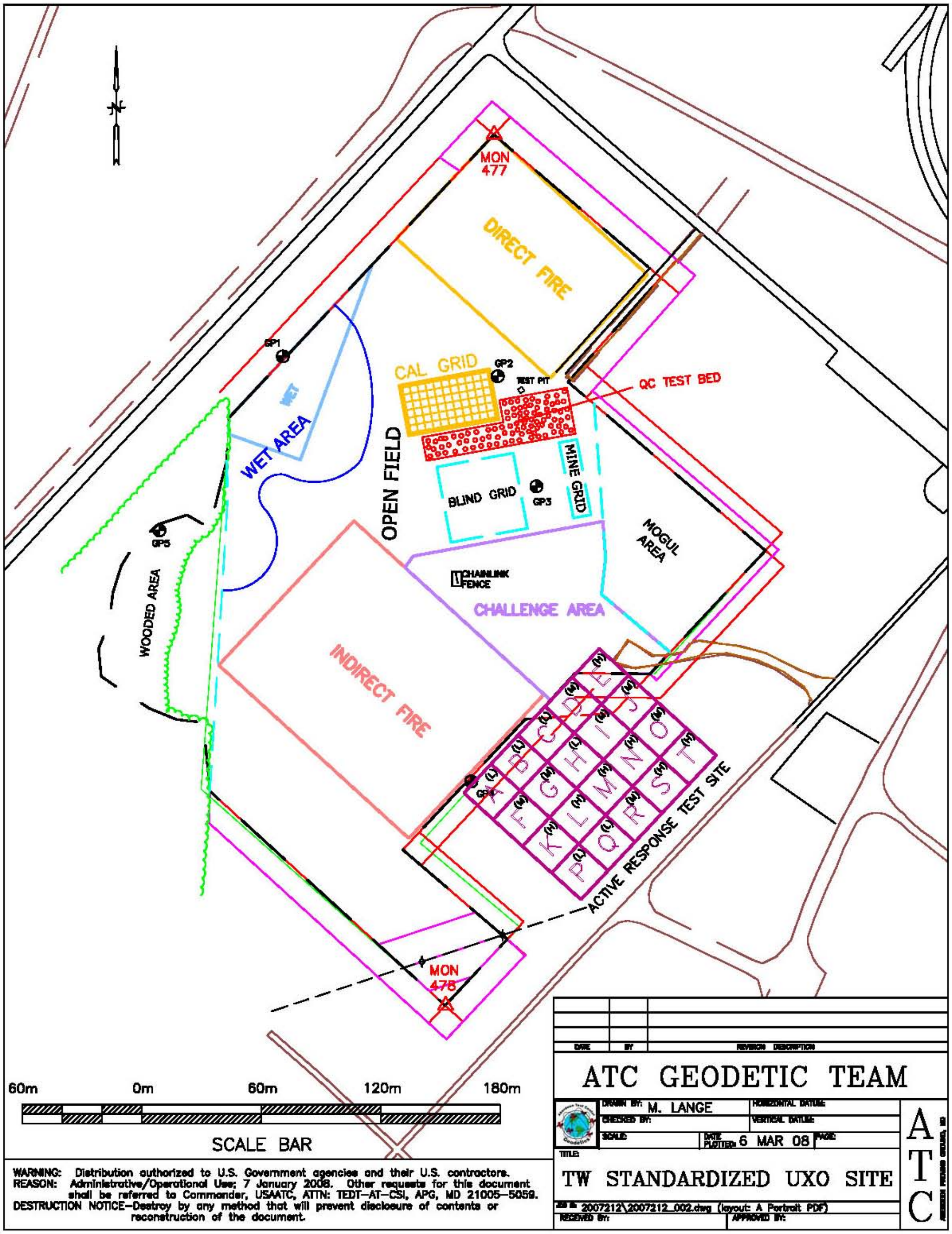


Figure 2. Test site layout.

2.2.4 STANDARD AND NONSTANDARD INERT MUNITIONS TARGETS

The standard and nonstandard munitions items emplaced in the test areas are presented in Table 2. Standardized targets are members of a set of specific munitions items that have identical properties to all other items in the set (caliber, configuration, size, weight, aspect ratio, material, filler, magnetic remanence, and nomenclature). Nonstandard targets are inert munitions items having properties that differ from those in the set of standardized items.

TABLE 2. INERT MUNITIONS TARGETS

Item	Munition Type	Calibration Lanes	Blind Grid	Open Field Direct Fire	Open Field Indirect Fire	Open Field Legacy	Moguls	Woods
20-mm Projectile M55	S	X				X	X	X
25-mm Projectile M794	S	X	X	X				
37-mm Projectile M47	S	X	X	X				
40-mm Projectile MKII Bodies	S	X				X	X	X
BDU-28 Submunition	S	X				X	X	X
BLU-26 Submunition	S	X				X	X	X
M42 Submunition	S	X				X	X	X
57-mm Projectile APC M86	S	X				X	X	X
60-mm Mortar M49A3	S	X	X		X			
2.75-in. Rocket M230	S	X				X	X	X
81-mm Mortar M374	S	X	X		X	X	X	X
105-mm HEAT Rounds M456	S					X	X	X
105-mm HEAT Round M490	S	X	X	X				
105-mm Projectile M60	S	X	X		X	X	X	X
155-mm Projectile M483A1	S	X				X	X	X
20-mm Projectile M55	NS					X	X	X
20-mm Projectile M97	NS					X	X	X
40-mm Projectile M813	NS					X	X	X
60-mm Mortar (JPG)	NS					X	X	X
60-mm Mortar M49	NS					X	X	X
2.75-in. Rocket M230	NS					X	X	X
2.75-in. Rocket XM229	NS					X	X	X
81-mm Mortar (JPG)	NS					X	X	X
81-mm Mortar M374	NS					X	X	X
105-mm Projectile M60	NS					X	X	X
155-mm Projectile M483A	NS					X	X	X

S = Standard munition.
NS = Nonstandard munition.
JPG = Jefferson Proving Ground.
HEAT = high-explosive antitank.

2.3 ATC SURVEY COMMENTS

None.

SECTION 3. FIELD DATA

3.1 DATE OF FIELD ACTIVITIES (16 through 18 November 2009)

3.2 AREAS TESTED/NUMBER OF HOURS

Areas tested and total numbers of hours operated at each site are presented in Table 3.

TABLE 3. AREAS TESTED AND NUMBER OF HOURS

Area	Number of Hours
Calibration lanes	3.0
Blind grid	3.16
Open field	11.83
Woods	NA
Mogul	NA
Mine grid	NA

Note: Table 3 represents the total time spent in each area.

3.3 TEST CONDITIONS

3.3.1 Weather Conditions

An APG weather station located approximately 1 mile west of the test site was used to record average temperature and precipitation on a half-hour basis for each day of operation. The temperatures presented in Table 4 represent the average temperature during field operations from 0700 to 1700 hours, while precipitation data represent a daily total amount of rainfall. Hourly weather logs used to generate this summary are provided in Appendix B.

TABLE 4. TEMPERATURE/PRECIPITATION DATA SUMMARY

Date, 09	Average Temperature, °F	Total Daily Precipitation, in.
16 Nov	62.1	0.00
17 Nov	53.1	0.00
18 Nov	53.9	0.00

3.3.2 Field Conditions

SAIC surveyed the calibration grid, blind grid, and indirect fire area of the open field. Numerous puddles and wet areas from rain prior to testing were present in the indirect fire area.

3.3.3 Soil Moisture

Three soil probes were placed at various locations within the site to capture soil moisture data: blind grid, calibration, open field, and wooded areas. Measurements were collected in percent moisture and were taken twice daily (morning and afternoon) from five different soil depths (1 to 6 in., 6 to 12 in., 12 to 24 in., 24 to 36 in., and 36 to 48 in.) from each probe. Soil moisture logs are provided in Appendix C.

3.4 FIELD ACTIVITIES

3.4.1 Setup/Mobilization

These activities included initial mobilization and daily equipment preparation and breakdown. A two-person crew took 2 hours and 35 minutes to perform the initial setup and mobilization. A total of 4 hours of equipment preparation was accrued, and end of day equipment breakdown totalled 1 hour and 5 minutes.

3.4.2 Calibration

SAIC spent a total of 3 hours in the calibration lanes, of which 2 hours and 25 minutes were spent collecting data.

3.4.3 Downtime Occasions

Occasions of downtime are grouped into five categories: equipment/data checks or equipment maintenance, equipment failure and repair, weather, demonstration site issues, or breaks/lunch. All downtime is included for the purposes of calculating labor requirements (section 5) except for downtime due to demonstration site issues. Demonstration site issues, while noted in the daily log, are considered nonchargeable downtime for the purposes of calculating labor costs and are not discussed. Breaks and lunches are discussed in this section and billed to the total site survey area.

3.4.3.1 Equipment/data checks, maintenance. Equipment data checks and maintenance activities accounted for 4 hours and 25 minutes of site usage time. These activities included changing out batteries and performing routine data checks to ensure the data were being properly recorded/collected. SAIC spent 55 minutes for breaks and lunches.

3.4.3.2 Equipment failure or repair. No equipment failure or repair occurred during this survey.

3.4.3.3 Weather. No weather delays occurred during the survey.

3.4.4 Data Collection

TABLE 5. TOTAL TIME
SAIC, SPENT PER AREA

Area	Time, hr/min
Blind grid	3 hours, 10 minutes
Open field	NA
Legacy	NA
Direct fire	NA
Indirect fire	11 hours, 50 minutes
Challenge	NA
Wooded	NA
Mine Grid	NA
Moguls	NA

Note: Table 5 represents the total time spent in each area collecting data.

3.4.5 Demobilization

The SAIC survey crew conducted a demonstration of the calibration grid and indirect fire. Demobilization occurred on 18 November 2009. On that day, it took the crew 1 hour and 35 minutes to break down and pack up their equipment.

3.5 PROCESSING TIME

SAIC submitted the raw data from the demonstration activities on the last day of the demonstration, as required. The scoring submittal data were provided in December 2009.

3.6 DEMONSTRATOR'S FIELD PERSONNEL

Rob Seigel
Chris Gibson

3.7 DEMONSTRATOR'S FIELD SURVEYING METHOD

SAIC collected the data in a linear fashion, using a line spacing of 1 meter.

3.8 SUMMARY OF DAILY LOGS

Daily logs capture all field activities during this demonstration and are provided in Appendix D.

SECTION 4. TECHNICAL PERFORMANCE RESULTS

4.1 ROC CURVES USING ALL MUNITIONS CATEGORIES

The probability of detection for the response stage (P_d^{res}) and the discrimination stage (P_d^{disc}) versus their respective probability of clutter detection or probability of false positive within each area are shown in Figures 3 through 8. The probabilities plotted against their respective background alarm rate within each area are shown in Figures 9 through 14. Both figures use horizontal lines to illustrate the performance of the demonstrator at two demonstrator-specified points: at the system noise level for the response stage, representing the point below which targets are not considered detectable, and at the demonstrator's recommended threshold level for the discrimination stage, defining the subset of targets the demonstrator would recommend digging based on discrimination. Note that all points have been rounded to protect the GT.

Not covered

Figure 3. SCEMP/pushcart blind grid probability of detection for response and discrimination stages versus their respective probability of false positive.

Not covered

Figure 4. SCEMP/pushcart open field (direct fire) probability of detection for response and discrimination stages versus their respective probability of false positive.

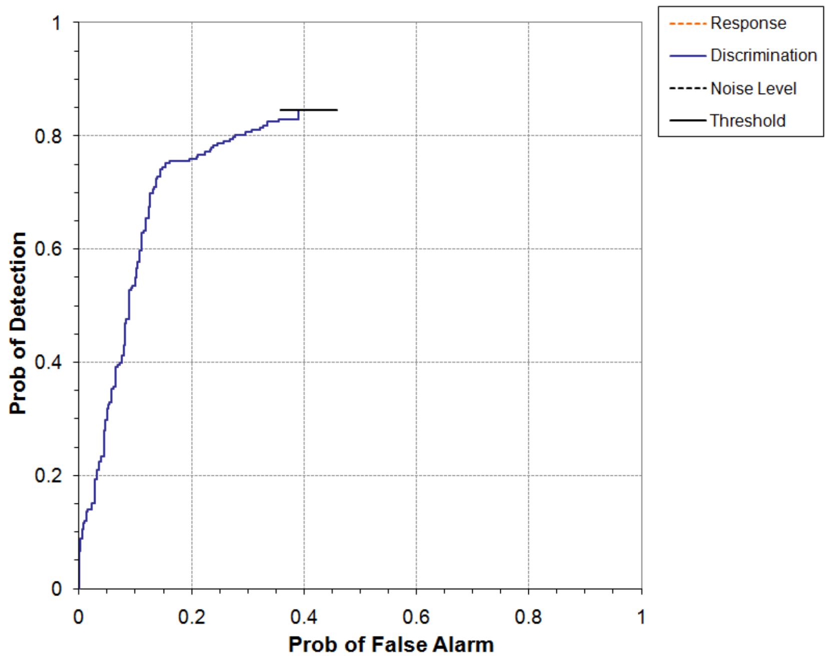


Figure 5. SCEMP/pushcart open field (indirect fire) probability of detection for response and discrimination stages versus their respective probability of false positive.

Not covered

Figure 6. SCEMP/pushcart open field (legacy) probability of detection for response and discrimination stages versus their respective probability of false positive.

Not covered

Figure 7. SCEMP/pushcart wooded probability of detection for response and discrimination stages versus their respective probability of false positive.

Not covered

Figure 8. SCEMP/pushcart mogul probability of detection for response and discrimination stages versus their respective probability of false positive.

Not covered

Figure 9. SCEMP/pushcart blind grid probability of detection for response and discrimination stages versus their respective probability of background alarm.

Not covered

Figure 10. SCEMP/pushcart open field (direct fire) probability of detection for response and discrimination stages versus their respective background alarm rate.

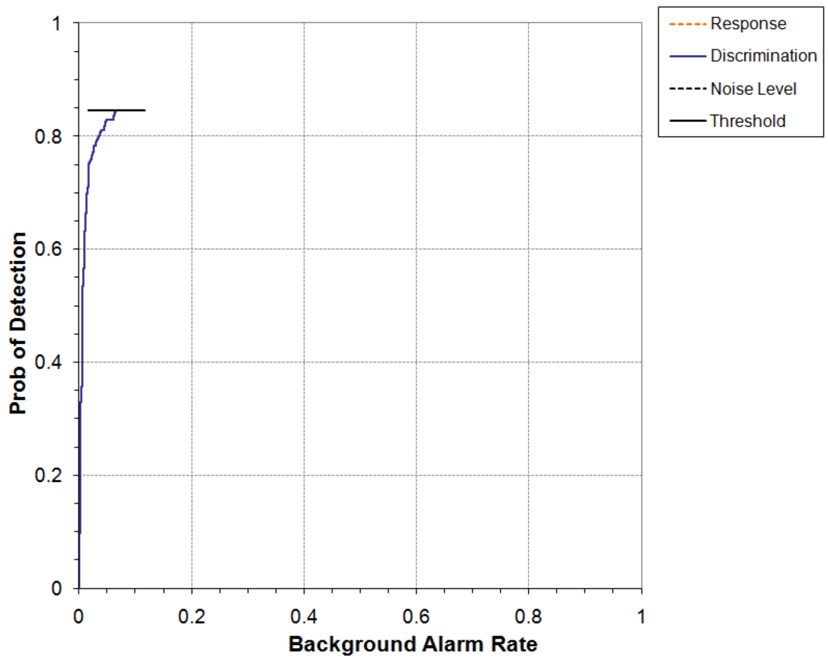


Figure 11. SCEMP/pushcart open field (indirect fire) probability of detection for response and discrimination stages versus their respective background alarm rate.

Not covered

Figure 12. SCEMP/pushcart open field (legacy) probability of detection for response and discrimination stages versus their respective background alarm rate.

Not covered

Figure 13. SCEMP/pushcart wooded probability of detection for response and discrimination stages versus their respective background alarm rate.

Not covered

Figure 14. SCEMP/pushcart mogul probability of detection for response and discrimination stages versus their respective background alarm rate.

4.2 PERFORMANCE SUMMARIES

Results for each of the testing areas are presented in Tables 6a through 6f (for labor requirements, see section 5). The response stage results are derived from the list of anomalies above the demonstrator-provided noise level. The results for the discrimination stage are derived from the demonstrator's recommended threshold for optimizing munitions related cleanup by minimizing false alarm digs and maximizing munitions recovery. The lower and upper 90-percent confidence limits on P_d, P_{cd}, and P_{fp} were calculated assuming that the number of detections and false positives are binomially distributed random variables.

TABLE 6a. BLIND GRID TEST AREA RESULTS (not covered)

	Response Stage				**Discrimination Stage**			
Munitions[a] Scores	***P_d^{res}: by type***				***P_d^{disc}: by type***			
	All Types	**105-mm**	**81/60-mm**	**37/25-mm**	**All Types**	**105-mm**	**81/60-mm**	**37/25-mm**
	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --
				By Depth[b]				
0 to 4D	--	--	--	--	--	--	--	--
4D to 8D	--	--	--	--	--	--	--	--
8D to 12D	--	--	--	--	--	--	--	--
Clutter Scores	***P_{cd}***				***P_{fp}***			
				By Mass				
By Depth[b]	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**
All Depth	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --
0 to 0.15 m	--	--	--	--	--	--	--	--
0.15 to 0.3 m	--	--	--	--	--	--	--	--
0.3 to 0.6 m	--	--	--	--	--	--	--	--
				Background Alarm Rates				
	P_{ba}^{res}: --				**P_{ba}^{disc}: --**			

[a]In cells with offset data entries, the numbers to the left are the result and the two numbers to the right are an upper and lower 90-percent confidence interval for an assumed binomial distribution.

[b]All depths are measured to the center of the object.

TABLE 6b. OPEN FIELD DIRECT FIRE TEST AREA RESULTS (not covered)

	Response Stage				Discrimination Stage			
Munitions[a] Scores	***P_d^{res}: by type***				***P_d^{disc}: by type***			
	All Types	**105-mm**	**37-mm**	**25-mm**	**All Types**	**105-mm**	**37-mm**	**25-mm**
	--	--	--	--	--	--	--	--
	--	--	--	--	--	--	--	--
	--	--	--	--	--	--	--	--
By Density								
High	--	--	--	--	--	--	--	--
Medium	--	--	--	--	--	--	--	--
Low	--	--	--	--	--	--	--	--
By Depth[b]								
0 to 4D	--	--	--	--	--	--	--	--
4D to 8D	--	--	--	--	--	--	--	--
8D to 12D	--	--	--	--	--	--	--	--
Clutter Scores	***P_{cd}***				***P_{fp}***			
By Mass								
By Depth[b]	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**
All Depth	--	--	--	--	--	--	--	--
	--	--	--	--	--	--	--	--
	--	--	--	--	--	--	--	--
0 to 0.15 m	--	--	--	--	--	--	--	--
0.15 to 0.3 m	--	--	--	--	--	--	--	--
0.3 to 0.6 m	--	--	--	--	--	--	--	--
Background Alarm Rates								
	BAR^{res}: --				**BAR^{disc}: --**			
Groups								
Found	--				--			
Identified	--				--			
Coverage	--				--			

[a]In cells with offset data entries, the numbers to the left are the result and the two numbers to the right are an upper and lower 90-percent confidence interval for an assumed binomial distribution.
[b]All depths are measured to the center of the object.

TABLE 6c. OPEN FIELD INDIRECT FIRE TEST AREA RESULTS

	Response Stage				Discrimination Stage			
Munitions[a] Scores	***P_d^{res}: by type***				***P_d^{disc}: by type***			
	All Types	**105-mm**	**81-mm**	**60-mm**	**All Types**	**105-mm**	**81-mm**	**60-mm**
	0.85 (0.87, 0.81)	0.89 (0.93, 0.83)	0.88 (0.92, 0.82)	0.77 (0.83, 0.71)	0.85 (0.87, 0.81)	0.89 (0.93, 0.83)	0.88 (0.92, 0.82)	0.77 (0.83, 0.71)
By Density								
High	0.73	0.85	0.82	0.52	0.73	0.85	0.82	0.52
Medium	0.84	0.87	0.86	0.79	0.84	0.87	0.86	0.79
Low	0.94	0.94	0.94	0.94	0.94	0.94	0.94	0.94
By Depth[b]								
0 to 4D	0.91	0.96	1.00	0.81	0.91	0.96	1.00	0.81
4D to 8D	0.79	0.84	0.80	0.64	0.79	0.84	0.80	0.64
8D to 12D	0.68	0.25	0.78	0.75	0.68	0.25	0.78	0.75
Clutter Scores	***P_{cd}***				***P_{fp}***			
By Mass								
By Depth[b]	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**	**All Mass**	**0 to 0.25 kg**	**>0.25 to 1 kg**	**>1 to 8 kg**
All Depth	0.41 (0.44, 0.38)	0.20	0.61	0.74	0.41 (0.44, 0.38)	0.20	0.61	0.74
0 to 0.15 m	0.39	0.20	0.63	0.75	0.39	0.20	0.63	0.75
0.15 to 0.3 m	0.45	0.25	0.44	0.65	0.45	0.25	0.44	0.65
0.3 to 0.6 m	0.75	0.00	0.67	1.00	0.75	0.00	0.67	1.00
Background Alarm Rates								
	BAR^{res}: 0.07				**BAR^{disc}:** 0.07			
Groups								
Found	0.90				0.90			
Identified	0.00				0.00			
Coverage	0.44				0.44			

[a]In cells with offset data entries, the numbers to the left are the result and the two numbers to the right are an upper and lower 90-percent confidence interval for an assumed binomial distribution.
[b]All depths are measured to the center of the object.

TABLE 6d. OPEN FIELD LEGACY TEST AREA RESULTS (not covered)

	Response Stage				**Discrimination Stage**			
Munitions[a] Scores	P_d^{res}: *by type*				P_d^{disc}: *by type*			
	All Types	**Small**	**Medium**	**Large**	**All Types**	**Small**	**Medium**	**Large**
	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --
By Depth[b]								
0 to 4D	--	--	--	--	--	--	--	--
4D to 8D	--	--	--	--	--	--	--	--
8D to 12D	--	--	--	--	--	--	--	--
> 12D	--	--	--	--	--	--	--	--

Clutter Scores	P_{cd}					P_{fp}				
	By Mass									
By Depth[b]	**All Mass**	0 to 0.25 kg	**>0.25 to 1 kg**	**>1 to 10 kg**	**> 10 kg**	**All Mass**	0 to 0.25 kg	**>0.25 to 1 kg**	**>1 to 8 kg**	**< 10kg**
All Depth	-- -- --	--	--	--	--	-- -- --	--	--	--	--
0 to 0.15 m	--	--	--	--	--	--	--	--	--	--
0.15 to 0.3 m	--	--	--	--	--	--	--	--	--	--
0.3 to 0.6 m	--	--	--	--	--	--	--	--	--	--
> 0.6 m	--	--	--	--	--	--	--	--	--	--

	Background Alarm Rates							
	BAR[res]:				**BAR[disc]:**			
	Groups							
Found	--				--			
Identified	--				--			
Coverage	--				--			

[a]In cells with offset data entries, the numbers to the left are the result and the two numbers to the right are an upper and lower 90-percent confidence interval for an assumed binomial distribution.
[b]All depths are measured to the center of the object.

TABLE 6e. WOODED TEST AREA RESULTS (not covered)

Response Stage					Discrimination Stage				
Munitions[a] Scores	P_d^{res}: *by type*				P_d^{disc}: *by type*				
	All Types	**Small**	**Medium**	**Large**	**All Types**	**Small**	**Medium**	**Large**	
	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	
By Depth[b]									
0 to 4D	--	--	--	--	--	--	--	--	
4D to 8D	--	--	--	--	--	--	--	--	
8D to 12D	--	--	--	--	--	--	--	--	
> 12D	--	--	--	--	--	--	--	--	

Clutter Scores	P_{cd}					P_{fp}				
By Mass										
By Depth[b]	**All Mass**	0 to 0.25 kg	**>0.25 to 1 kg**	**>1 to 10 kg**	**> 10 kg**	**All Mass**	0 to 0.25 kg	**>0.25 to 1 kg**	**>1 to 8 kg**	**< 10kg**
All Depth	-- -- --	--	--	--	--	-- -- --	--	--	--	--
0 to 0.15 m	--	--	--	--	--	--	--	--	--	--
0.15 to 0.3 m	--	--	--	--	--	--	--	--	--	--
0.3 to 0.6 m	--	--	--	--	--	--	--	--	--	--
> 0.6 m	--	--	--	--	--	--	--	--	--	--
Background Alarm Rates										
	BAR^{res}**:**					**BAR**^{disc}**:**				
Groups										
Found	--					--				
Identified	--					--				
Coverage	--					--				

[a]In cells with offset data entries, the numbers to the left are the result and the two numbers to the right are an upper and lower 90-percent confidence interval for an assumed binomial distribution.
[b]All depths are measured to the center of the object.

TABLE 6f. MOGUL TEST AREA RESULTS (not covered)

	Response Stage				**Discrimination Stage**			
Munitions[a] Scores	P_d^{res}: *by type*				P_d^{disc}: *by type*			
	All Types	**Small**	**Medium**	**Large**	**All Types**	**Small**	**Medium**	**Large**
	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --	-- -- --
By Depth[b]								
0 to 4D	--	--	--	--	--	--	--	--
4D to 8D	--	--	--	--	--	--	--	--
8D to 12D	--	--	--	--	--	--	--	--
> 12D	--	--	--	--	--	--	--	--

Clutter Scores	P_{cd}					P_{fp}				
	By Mass									
By Depth[b]	**All Mass**	0 to 0.25 kg	**>0.25 to 1 kg**	**>1 to 10 kg**	**> 10 kg**	**All Mass**	0 to 0.25 kg	**>0.25 to 1 kg**	**>1 to 8 kg**	**< 10kg**
All Depth	-- -- --	--	--	--	--	-- -- --	--	--	--	--
0 to 0.15 m	--	--	--	--	--	--	--	--	--	--
0.15 to 0.3 m	--	--	--	--	--	--	--	--	--	--
0.3 to 0.6 m	--	--	--	--	--	--	--	--	--	--
> 0.6 m	--	--	--	--	--	--	--	--	--	--

	Background Alarm Rates	
	BAR^{res}:	**BAR^{disc}:**

	Groups							
Found	--				--			
Identified	--				--			
Coverage	--				--			

[a]In cells with offset data entries, the numbers to the left are the result and the two numbers to the right are an upper and lower 90-percent confidence interval for an assumed binomial distribution.
[b]All depths are measured to the center of the object.

4.3 EFFICIENCY, REJECTION RATES, AND TYPE CLASSIFICATION

Efficiency and rejection rates are calculated to quantify the discrimination ability at specific points of interest on the ROC curve: (1) at the point where no decrease in P_d is suffered (i.e., the efficiency is by definition equal to one) and (2) at the operator selected threshold. These values are presented in Tables 7a through 7f.

TABLE 7a. BLIND GRID EFFICIENCY AND REJECTION RATES (not covered)

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	--	--	--
With No Loss of P_d	--	--	--

TABLE 7b. OPEN FIELD (DIRECT) EFFICIENCY AND REJECTION RATES (not covered)

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	--	--	--
With No Loss of P_d	--	--	--

TABLE 7c. OPEN FIELD (INDIRECT) EFFICIENCY AND REJECTION RATES

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	1.00	0.00	0.00
With No Loss of P_d	1.00	0.05	0.05

TABLE 7d. OPEN FIELD (LEGACY) EFFICIENCY AND REJECTION RATES (not covered)

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	--	--	--
With No Loss of P_d	--	--	--

TABLE 7e. WOODED EFFICIENCY AND REJECTION RATES (not covered)

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	--	--	--
With No Loss of P_d	--	--	--

TABLE 7f. MOGUL EFFICIENCY AND REJECTION RATES (not covered)

	Efficiency (E)	False Positive Rejection Rate	Background Alarm Rejection Rate
At Operating Point	--	--	--
With No Loss of P_d	--	--	--

At the demonstrator's recommended setting, the munitions items that were detected and correctly discriminated were further scored on whether their correct type could be identified (table 8a through 8f). Correct type examples include 20-mm projectile, 105-mm HEAT projectile, and 2.75-inch Rocket. A list of the standard type declaration required for each munitions item was provided to demonstrators prior to testing. The standard types for the three example items are 20-mmP, 105H, and 2.75-inch.

TABLE 8a. BLIND GRID CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS MUNITIONS (not covered)

Size	Percentage Correct
25mm	--
37mm	--
60mm	--
81mm	--
105mm	--
105 artillery	--
Overall	--

TABLE 8b. OPEN FIELD DIRECT FIRE CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS MUNITIONS (not covered)

Size	Percentage Correct
25mm	--
37mm	--
105mm	--
Overall	--

TABLE 8c. OPEN FIELD INDIRECT FIRE CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS MUNITIONS

Size	Percentage Correct
60mm	[a]NA
81mm	[a]NA
105mm	[a]NA
Overall	[a]NA

[a]Vendor covered the area but did not identify the item size.

TABLE 8d. OPEN FIELD LEGACY CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS MUNITIONS (not covered)

Size	Percentage Correct
Small	--
Medium	--
Large	--
Overall	--

TABLE 8e. WOODED CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS MUNITIONS (not covered)

Size	Percentage Correct
Small	--
Medium	--
Large	--
Overall	--

TABLE 8f. MOGUL CORRECT TYPE CLASSIFICATION OF TARGETS CORRECTLY DISCRIMINATED AS MUNITIONS (not covered)

Size	Percentage Correct
Small	--
Medium	--
Large	--
Overall	--

4.4 LOCATION ACCURACY

The mean location error and standard deviations appear in Tables 9a through 9f. These calculations are based on average missed distance for munitions correctly identified during the response stage. Depths are measured from the center of the munitions to the surface. For the blind grid, only depth errors are calculated because (X, Y) positions are known to be the centers of the grid square.

TABLE 9a. BLIND GRID MEAN LOCATION ERROR AND STANDARD DEVIATION(not covered)

	Mean	Standard Deviation
Northing	--	--
Easting	--	--
Depth	--	--

TABLE 9b. OPEN FIELD DIRECT FIRE MEAN LOCATION ERROR AND STANDARD DEVIATION (not covered)

	Mean	Standard Deviation
Northing	--	--
Easting	--	--
Depth	--	--

TABLE 9c. OPEN FIELD INDIRECT FIRE MEAN LOCATION ERROR AND STANDARD DEVIATION

	Mean	Standard Deviation
Northing	0.01	0.17
Easting	0.00	0.17
Depth	[a]NA	[a]NA

[a]Vendor did not identify the time depth.

TABLE 9d. OPEN FIELD LEGACY MEAN LOCATION ERROR AND STANDARD DEVIATION (not covered)

	Mean	Standard Deviation
Northing	--	--
Easting	--	--
Depth	--	--

TABLE 9e. WOODED MEAN LOCATION ERROR AND STANDARD DEVIATION (not covered)

	Mean	Standard Deviation
Northing	--	--
Easting	--	--
Depth	--	--

TABLE 9f. MOGUL MEAN LOCATION ERROR AND STANDARD DEVIATION (not covered)

	Mean	Standard Deviation
Northing	--	--
Easting	--	--
Depth	--	--

SECTION 5. APPENDIXES

APPENDIX A. TERMS AND DEFINITIONS

GENERAL DEFINITIONS

Anomaly: Location of a system response deemed to warrant further investigation by the demonstrator for consideration as an emplaced munitions item.

Detection: An anomaly location that is within R_{halo} of an emplaced munitions item.

Military Munitions (MM): Specific categories of MM that may pose unique explosive safety risks, including UXO as defined in 10 USC 101(e)(5), DMM as defined in 10 USC 2710(e)(2) and/or munitions constituents (e.g., TNT, RDX) as defined in 10 USC 2710(e)(3) that are present in high enough concentrations to pose an explosive hazard.

Emplaced Munitions: A munitions item buried by the government at a specified location in the test site.

Emplaced Clutter: A clutter item (i.e., nonmunitions item) buried by the government at a specified location in the test site.

R_{halo}: A predetermined radius about an emplaced item (clutter or munitions) within which an anomaly identified by the demonstrator as being of interest is considered to be a detection of that item. For the purpose of this program, a circular halo 0.5 meters in radius is placed around the center of the object for all clutter and munitions items.

Small Munitions: Caliber of munitions less than or equal to 40 mm (includes 20-mm projectile, 25-mm projectile, 37-mm projectile, 40-mm projectile, submunitions BLU-26, BLU-63, and M42).

Medium Munitions: Caliber of munitions greater than 40 mm and less than or equal to 81 mm (includes 57-mm projectile, 60-mm mortar, 2.75-inch rocket, and 81-mm mortar).

Large Munitions: Caliber of munitions greater than 81 mm (includes 105-mm HEAT, 105-mm projectile, and 155-mm projectile).

Group: Two or more adjacent GT items with overlapping halos.

GT: Ground truth

Response Stage Noise Level: The level that represents the signal level below which anomalies are not considered detectable. Demonstrators are required to provide the recommended noise level for the blind grid test area.

Discrimination Stage Threshold: The demonstrator-selected threshold level that is expected to provide optimum performance of the system by retaining all detectable munitions and rejecting the maximum amount of clutter. This level defines the subset of anomalies the demonstrator would recommend digging based on discrimination.

Binomially Distributed Random Variable: A random variable of the type which has only two possible outcomes, say success and failure, is repeated for *n* independent trials with the probability *p* of success and the probability *1-p* of failure being the same for each trial. The number of successes *x* observed in the *n* trials is an estimate of *p* and is considered to be a binomially distributed random variable.

RESPONSE AND DISCRIMINATION STAGE DATA

The scoring of the demonstrator's performance is conducted in two stages: response stage and discrimination stage. For both stages, the probability of detection (P_d) and the false alarms are reported as receiver-operating characteristic (ROC) curves. False alarms are divided into those anomalies that correspond to emplaced clutter items, measuring the probability of clutter detection (P_{cd}) or probability of false positive (P_{fp}). Those that do not correspond to any known item are termed background alarms.

The response stage is a measure of whether the sensor can detect an object of interest. For a channel instrument, this value should be closely related to the amplitude of the signal. The demonstrator must report the response level (threshold) below which target responses are deemed insufficient to warrant further investigation. At this stage, minimal processing may be done. This includes filtering long- and short-scale variations, bias removal, and scaling. This processing should be detailed in the data submission.

For a multichannel instrument, the demonstrator must construct a quantity analogous to amplitude. The demonstrator should consider what combination of channels provides the best test for detecting any object that the sensor can detect. The average amplitude across a set of channels is an example of an acceptable response stage quantity. Other methods may be more appropriate for a given sensor. Again, minimal processing can be done, and the demonstrator should explain how this quantity was constructed in their data submission.

The discrimination stage evaluates the demonstrator's ability to correctly identify munitions as such, and to reject clutter. For the same locations as in the response stage anomaly list, the discrimination stage list contains the output of the algorithms applied in the discrimination-stage processing. This list is prioritized based on the demonstrator's determination that an anomaly location is likely to contain munitions. Thus, higher output values are indicative of higher confidence that a munitions item is present at the specified location. For electronic signal processing, priority ranking is based on algorithm output. For other systems, priority ranking is based on human judgment. The demonstrator also selects the threshold that the demonstrator believes will provide optimum system performance, (i.e., that retains all the detected munitions and rejects the maximum amount of clutter).

Note: The two lists provided by the demonstrator contain identical numbers of potential target locations. They differ only in the priority ranking of the declarations.

GROUP SCORING FACTORS

Based on configuration of the GT at the standardized sites and the defined scoring methodology, there exists munitions groups defined as having overlapping halos. In these cases, the following scoring logic is implemented (fig. A-1 through A-9):

a. Overall site scores (i.e., P_d) will consider only isolated munitions and clutter items.

b. GT items that have overlapping halos (both munitions and clutter) will form a group and groups may form chains.

c. Groups will have a complex halos composed of all the composite halos of all its GT items.

d. Groups will have three scoring factors: groups found groups identified and group coverage. Scores will be based on 1:1 matches of anomalies and GT.

(1) Groups Found (Found): the number of groups that have one or more GT items matched divided by the total number of groups. Demonstrators will be credited with detecting a group if any item within the group is matched to an anomaly in their list.

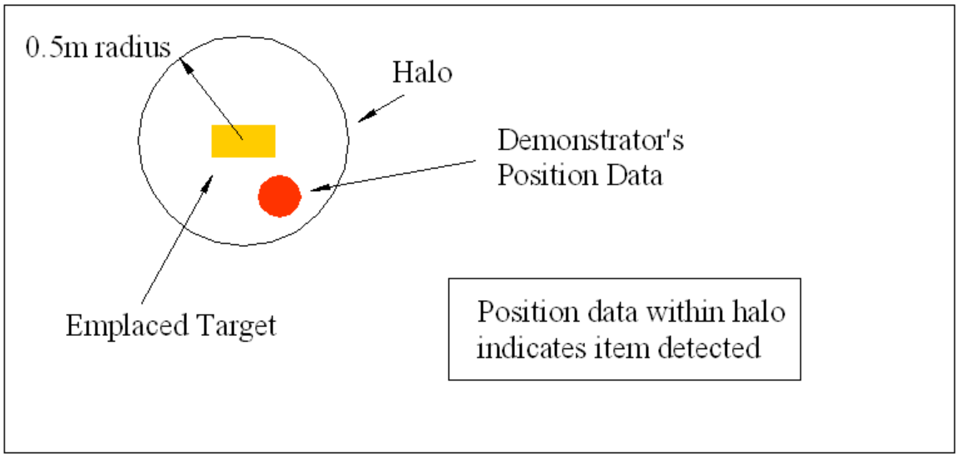
(2) Groups Identified (ID): the number of groups that have two or more GT items matched divided by the total number of groups. Demonstrators will be credited with identifying that a group is present if multiple items within the composite halo are matched to anomalies in their list.

(3) Group Coverage (Coverage): the number of GT items matched within groups divided by the total number of GT items within groups. This metric measures the demonstrator accuracy in determining the number of anomalies within a group. If five items are present and only two anomalies are matched, the demonstrator will score 0.4. If all five are matched the demonstrator will score 1.0.

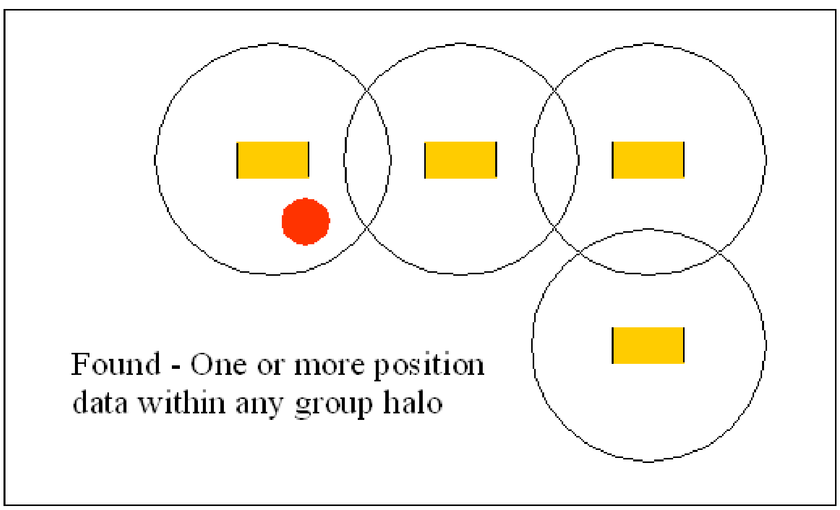
e. Location error will not be reported for groups.

f. Demonstrators will not be asked to call out groups in their scoring submissions. If multiple anomalies are indicated in a small area, the demonstrator will report all individual anomalies.

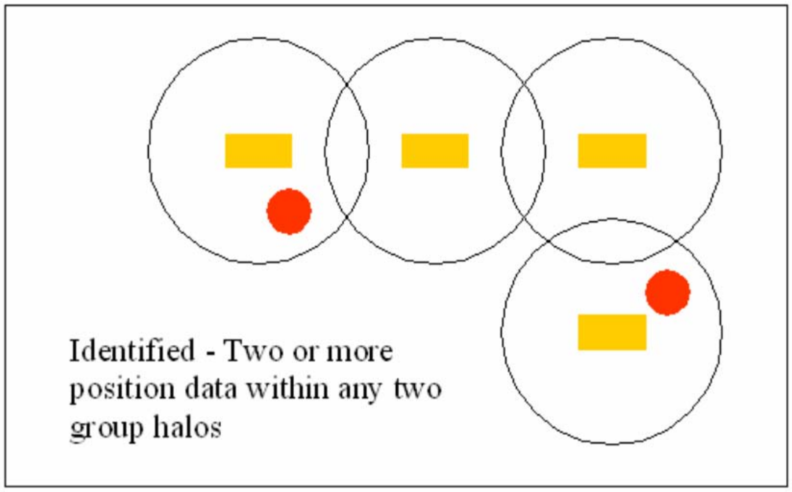
g. Excess alarms within a halo will be disregarded.



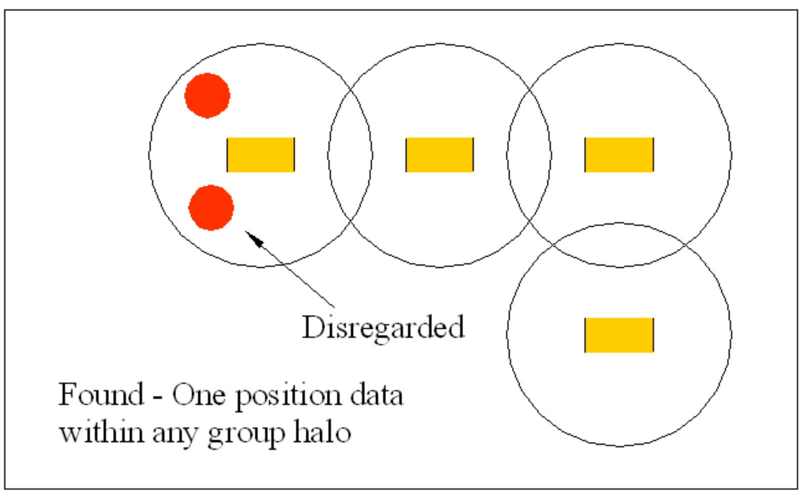
A-1. Example of detected item.



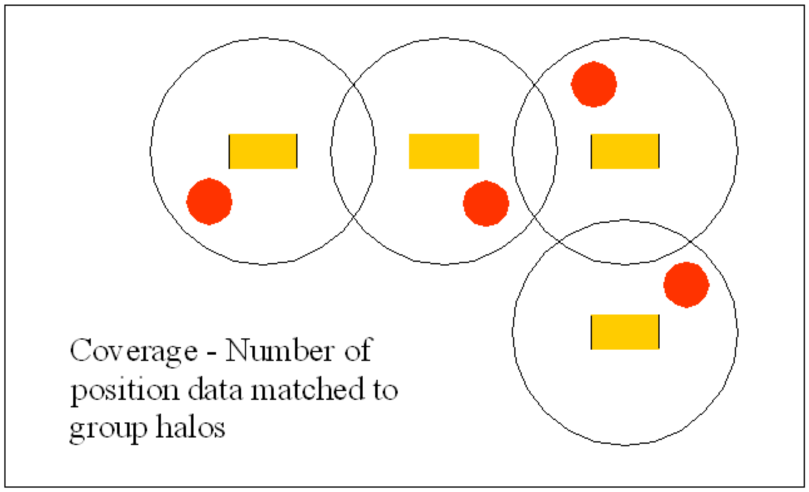
A-2. Example of group found (found).



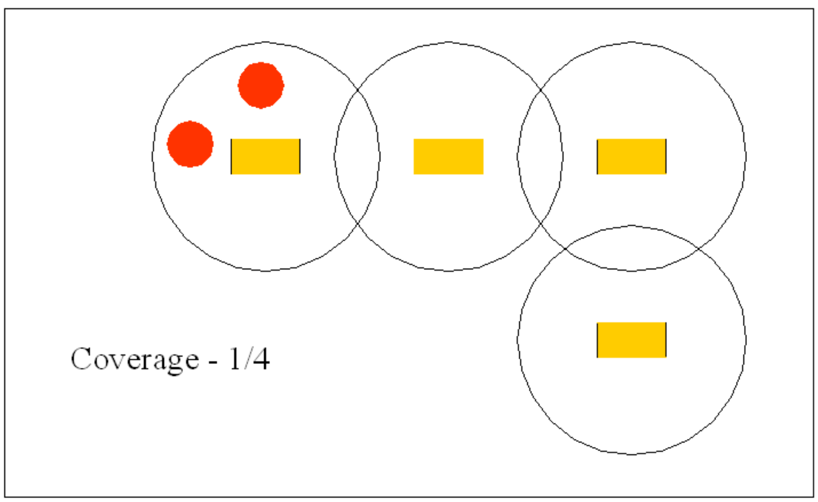
A-3. Example of group identified (ID).



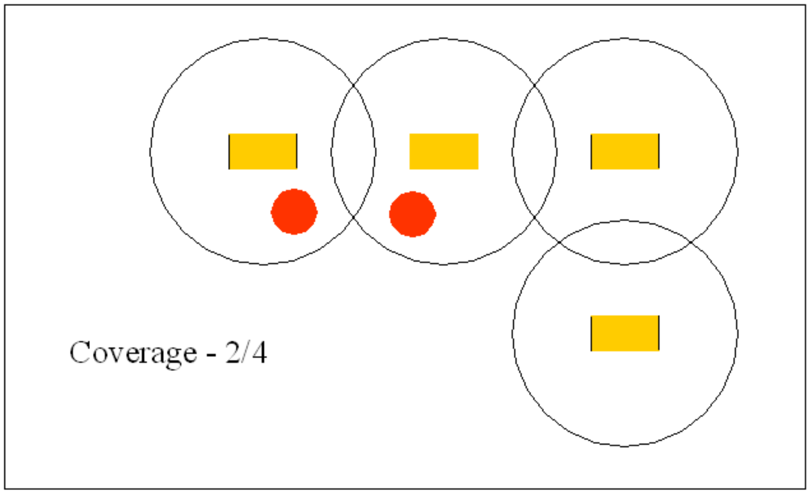
A-4. Example of excess alarms disregarded.



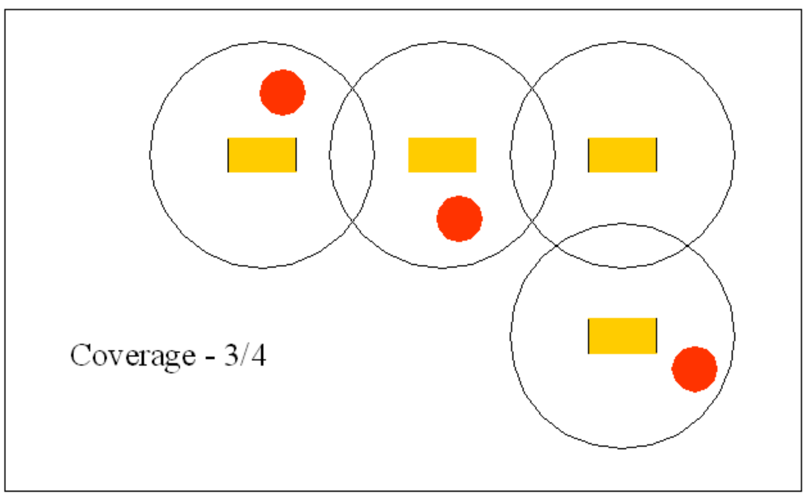
A-5. Example of a group.



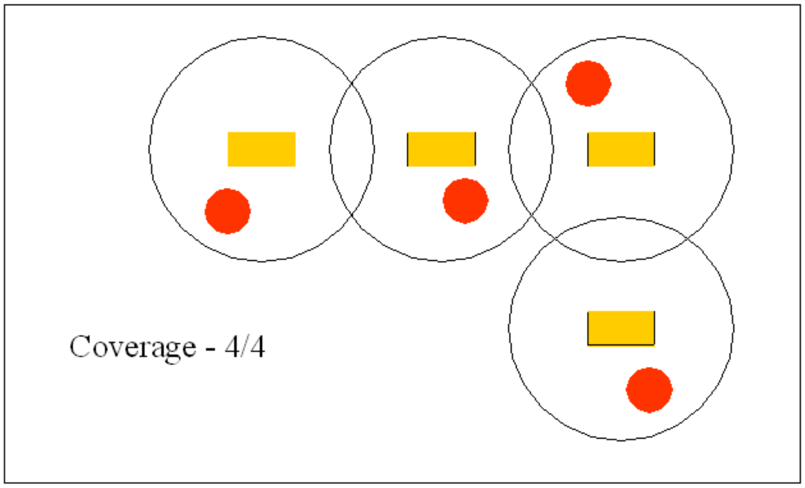
A-6. Example of group (1/4 = 0.25).



A-7. Example of group (2/4 = 0.5).



A-8. Example of group (3/4 = 0.75).



A-9. Example of group (4/4 = 1.0).

RESPONSE STAGE DEFINITIONS

Response Stage Probability of Detection (P_d^{res}): P_d^{res} = (No. of response-stage detections)/(No. of emplaced munitions in the test site).

Response Stage Clutter Detection (cd^{res}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Response Stage Probability of Clutter Detection (P_{cd}^{res}): P_{cd}^{res} = (No. of response-stage clutter detections)/(No. of emplaced clutter items).

Response Stage Background Alarm (ba^{res}): An anomaly in a blind grid cell that contains neither emplaced munitions nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced munitions or emplaced clutter item.

Response Stage Probability of Background Alarm (P_{ba}^{res}): Blind grid only: P_{ba}^{res} = (No. of response-stage background alarms)/(No. of empty grid locations).

Response Stage Background Alarm Rate (BAR^{res}): Open field any challenge area (including the direct and indirect firing sub areas) only: BAR^{res} = (No. of response-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{res}, P_{cd}^{res}, P_{ba}^{res}, and BAR^{res} are functions of t^{res}, the threshold applied to the response-stage signal strength. These quantities can therefore be written as $P_d^{res}(t^{res})$, $P_{cd}^{res}(t^{res})$, $P_{ba}^{res}(t^{res})$, and $BAR^{res}(t^{res})$.

DISCRIMINATION STAGE DEFINITIONS

Discrimination: The application of a signal processing algorithm or human judgment to sensor data to discriminate munitions from clutter. Discrimination should identify anomalies that the demonstrator has high confidence correspond to munitions, as well as those that the demonstrator has high confidence correspond to nonmunitions or background returns. The former should be ranked with highest priority and the latter with lowest.

Discrimination Stage Probability of Detection (P_d^{disc}): P_d^{disc} = (No. of discrimination-stage detections)/(No. of emplaced munitions in the test site).

Discrimination Stage False Positive (fp^{disc}): An anomaly location that is within R_{halo} of an emplaced clutter item.

Discrimination Stage Probability of False Positive (P_{fp}^{disc}): P_{fp}^{disc} = (No. of discrimination stage false positives)/(No. of emplaced clutter items).

Discrimination Stage Background Alarm (ba^{disc}): An anomaly in a blind grid cell that contains neither emplaced munitions nor an emplaced clutter item. An anomaly location in the open field or scenarios that is outside R_{halo} of any emplaced munitions or emplaced clutter item.

Discrimination Stage Probability of Background Alarm (P_{ba}^{disc}): P_{ba}^{disc} = (No. of discrimination-stage background alarms)/(No. of empty grid locations).

Discrimination Stage Background Alarm Rate (BAR^{disc}): BAR^{disc} = (No. of discrimination-stage background alarms)/(arbitrary constant).

Note that the quantities P_d^{disc}, P_{fp}^{disc}, P_{ba}^{disc}, and BAR^{disc} are functions of t^{disc}, the threshold applied to the discrimination-stage signal strength. These quantities can therefore be written as $P_d^{disc}(t^{disc})$, $P_{fp}^{disc}(t^{disc})$, $P_{ba}^{disc}(t^{disc})$, and $BAR^{disc}(t^{disc})$.

RECEIVER-OPERATING CHARACERISTIC (ROC) CURVES

ROC curves at both the response and discrimination stages can be constructed based on the above definitions. The ROC curves plot the relationship between P_d versus P_{cd} or P_{fp} and P_d versus BAR or P_{ba} as the threshold applied to the signal strength is varied from its minimum (t_{min}) to its maximum (t_{max}) value.[1] P_d versus P_{fp} and P_d versus BAR being combined into ROC curves is shown in Figure A-10. Note that the "res" and "disc" superscripts have been suppressed from all the variables for clarity.

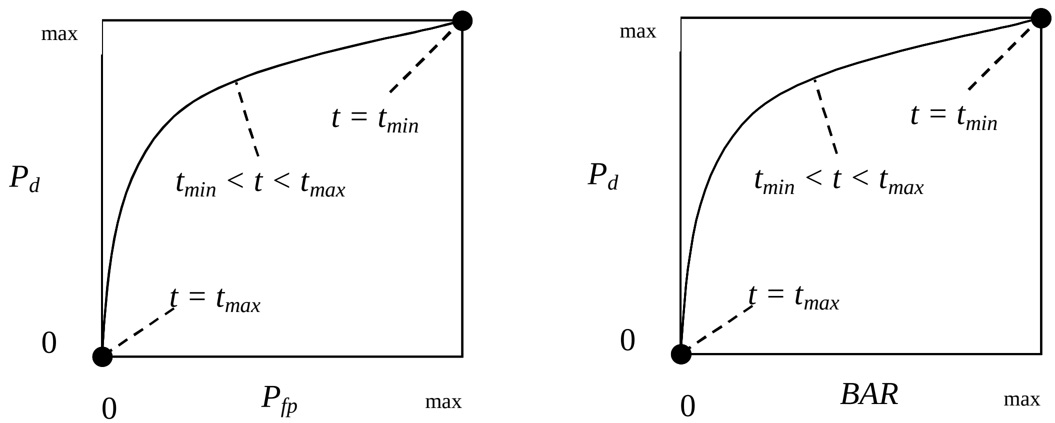


Figure A-10. ROC curves for open field testing. Each curve applies to both the response and discrimination stages.

METRICS TO CHARACTERIZE THE DISCRIMINATION STAGE

The demonstrator is also scored on efficiency and rejection ratio, which measure the effectiveness of the discrimination stage processing. The goal of discrimination is to retain the greatest number of munitions detections from the anomaly list while rejecting the maximum number of anomalies arising from nonmunitions items. The efficiency measures the fraction of detected munitions retained by the discrimination, while the rejection ratio measures the fraction of false alarms rejected. Both measures are defined relative to the entire response list, i.e., the maximum munitions detectable by the sensor and its accompanying clutter detection rate/false positive rate or background alarm rate.

[1]Strictly speaking, ROC curves plot the P_d versus P_{ba} over a predetermined and fixed number of detection opportunities (some of the opportunities are located over munitions and others are located over clutter or blank spots). In an open field scenario, each system suppresses its signal strength reports until some bare-minimum signal response is received by the system. Consequently, the open field ROC curves do not have information from low signal-output locations, and, furthermore, different contractors report their signals over a different set of locations on the ground. These ROC curves are thus not true to the strict definition of ROC curves as defined in textbooks on detection theory. Note, however, that the ROC curves obtained in the blind grid test sites are true ROC curves.

Efficiency (E): $E = P_d^{disc}(t^{disc})/P_d^{res}(t_{min}^{res})$: Measures (at a threshold of interest) the degree to which the maximum theoretical detection performance of the sensor system (as determined by the response stage tmin) is preserved after application of discrimination techniques. Efficiency is a number between 0 and 1. An efficiency of 1 implies that all of the munitions initially detected in the response stage were retained at the specified threshold in the discrimination stage, t^{disc}.

False Positive Rejection Rate (R_{fp}): $R_{fp} = 1 - [P_{fp}^{disc}(t^{disc})/P_{cd}^{res}(t_{min}^{res})]$: Measures (at a threshold of interest) the degree to which the sensor system's false positive performance is improved over the maximum false positive performance (as determined by the response stage tmin). The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all emplaced clutter initially detected in the response stage were correctly rejected at the specified threshold in the discrimination stage.

Background Alarm Rejection Rate (R_{ba}):

Blind grid: $R_{ba} = 1 - [P_{ba}^{disc}(t^{disc})/P_{ba}^{res}(t_{min}^{res})]$.
Open field: $R_{ba} = 1 - [BAR^{disc}(t^{disc})/BAR^{res}(t_{min}^{res})]$).

Measures the degree to which the discrimination stage correctly rejects background alarms initially detected in the response stage. The rejection rate is a number between 0 and 1. A rejection rate of 1 implies that all background alarms initially detected in the response stage were rejected at the specified threshold in the discrimination stage.

CHI-SQUARE COMPARISON

The Chi-square test for differences in probabilities (or 2 by 2 contingency table) is used to analyze two samples drawn from two different populations to see if both populations have the same or different proportions of elements in a certain category. More specifically, two random samples are drawn, one from each population, to test the null hypothesis that the probability of event A (some specified event) is the same for both populations.

The test statistic of the 2 by 2 contingency table is the Chi-square distribution with one degree of freedom. When an association between a more challenging terrain feature and relatively degraded performance is sought, a one-sided test is performed. A two-sided 2 by 2 contingency table is used in the Standardized UXO Technology Demonstration Site Program to compare performance between any two areas or subareas when the direction of degradation cannot be predetermined.

For a one-sided test, a significance level of 0.05 is used to set the critical decision limit. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, then the lower proportion tested will be considered significantly less than the greater one (degraded). If the test statistic calculated from the data is less than this value, then no degradation can be said to exist because of the terrain feature introduced.

For a two-sided test, a significance level of 0.10 is used to allow .05 on either side of the decision. It is a critical decision limit because if the test statistic calculated from the data exceeds this value, then the two proportions tested will be considered significantly different. If the test statistic calculated from the data is less than this value, then the two proportions tested will be considered not significantly different.

An exception must be applied when either a 0 or 100 percent success rate occurs in the sample data. The Chi-square test cannot be used in these instances. Instead, Fischer's test is used, and the critical decision limit for one-sided tests is the chosen significance level, which in this case is 0.05. With Fischer's test, if the test statistic is less than the critical value, then the proportions are considered to be significantly different.

An example follows that illustrates Standardized UXO Technology Demonstration Site blind grid results compared to those from the open field legacy. It should be noted that a significant result does not prove a cause-and-effect relationship exists between the two populations of interest; however, it does serve as a tool to indicate that one data set has experienced a degradation or change in system performance at a large enough level than can be accounted for merely by chance or random variation. Note also that a result that is not significant indicates that there is not enough evidence to declare that anything more than chance or random variation within the same population is at work between the two data sets being compared.

Demonstrator X achieves the following overall results after surveying the blind grid and open field (legacy) using the same system (results indicate the number of munitions detected divided by the number of munitions emplaced):

	Blind grid	Open field
P_d^{res}	100/100 = 1.0	8/10 = .80

P_d^{res}: BLIND GRID versus OPEN FIELD (legacy). Using the example data above to compare probabilities of detection in the response stage, all 100 munitions out of 100 emplaced munitions items were detected in the blind grid while 8 munitions out of 10 emplaced were detected in the open field. Fischer's test must be used since a 100 percent success rate occurs in the data. Fischer's test uses the four input values to calculate a test statistic of 0.0075 that is compared against the critical value of 0.05. Since the test statistic is less than the critical value, the smaller response stage detection rate (0.80) is considered to be significantly less at the 0.05 level of significance. While a significant result does not prove a cause-and-effect relationship exists between the change in survey area and degradation in performance, it does indicate that the detection ability of demonstrator X's system seems to have been degraded in the open field relative to results from the blind grid using the same system. This is an example of a one-sided Chi-squared test.

APPENDIX B. DAILY WEATHER LOGS

Date, 09	Time, [a]EST	Avg. Temp, °F	Total Precip., in.
16 Nov	0700	58.1	0.00
	0800	58.3	0.00
	0900	60.1	0.00
	1000	61.3	0.00
	1100	62.8	0.00
	1200	63.7	0.00
	1300	64.6	0.00
	1400	64.6	0.00
	1500	64.8	0.00
	1600	63.7	0.00
	1700	60.8	0.00
17 Nov	0700	45.1	0.00
	0800	46.0	0.00
	0900	49.3	0.00
	1000	51.4	0.00
	1100	53.4	0.00
	1200	54.9	0.00
	1300	55.9	0.00
	1400	57.4	0.00
	1500	58.3	0.00
	1600	57.6	0.00
	1700	55.0	0.00
18 Nov	0700	46.6	0.00
	0800	46.9	0.00
	0900	49.1	0.00
	1000	51.6	0.00
	1100	53.4	0.00
	1200	54.9	0.00
	1300	57.0	0.00
	1400	58.3	0.00
	1500	58.8	0.00
	1600	58.6	0.00
	1700	57.2	0.00

[a]Eastern Standard Time

APPENDIX C. SOIL MOISTURE

Date: 16 Nov 09			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Wet area	0 to 6	--	--
	6 to 12	--	--
	12 to 24	--	--
	24 to 36	--	--
	36 to 48	--	--
Wooded area	0 to 6	--	--
	6 to 12	--	--
	12 to 24	--	--
	24 to 36	--	--
	36 to 48	--	--
Open area	0 to 6	--	--
	6 to 12	--	--
	12 to 24	--	--
	24 to 36	--	--
	36 to 48	--	--
Calibration lanes	0 to 6	--	25.2
	6 to 12	--	28.6
	12 to 24	--	29.2
	24 to 36	--	31.2
	36 to 48	--	38.6
Blind grid/moguls	0 to 6	--	11.8
	6 to 12	--	26.3
	12 to 24	--	31.2
	24 to 36	--	35.8
	36 to 48	--	38.3

Date: 17 Nov 09			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Wet area	0 to 6	--	--
	6 to 12	--	--
	12 to 24	--	--
	24 to 36	--	--
	36 to 48	--	--
Wooded area	0 to 6	--	--
	6 to 12	--	--
	12 to 24	--	--
	24 to 36	--	--
	36 to 48	--	--
Open area	0 to 6	25.3	25.2
	6 to 12	28.7	28.6
	12 to 24	31.9	32.0
	24 to 36	40.5	40.3
	36 to 48	42.4	42.6
Calibration lanes	0 to 6	--	25.1
	6 to 12	--	28.4
	12 to 24	--	29.0
	24 to 36	--	31.5
	36 to 48	--	38.7
Blind grid/moguls	0 to 6	11.7	--
	6 to 12	26.1	--
	12 to 24	31.0	--
	24 to 36	35.8	--
	36 to 48	38.5	--

Date: 18 Nov 09			
Probe Location	**Layer, in.**	**AM Reading, %**	**PM Reading, %**
Wet area	0 to 6	--	--
	6 to 12	--	--
	12 to 24	--	--
	24 to 36	--	--
	36 to 48	--	--
Wooded area	0 to 6	--	--
	6 to 12	--	--
	12 to 24	--	--
	24 to 36	--	--
	36 to 48	--	--
Open area	0 to 6	25.0	24.8
	6 to 12	28.4	28.3
	12 to 24	31.8	31.9
	24 to 36	40.1	40.0
	36 to 48	42.5	42.4
Calibration lanes	0 to 6	--	24.8
	6 to 12	--	28.2
	12 to 24	--	28.7
	24 to 36	--	31.6
	36 to 48	--	38.9
Blind grid/moguls	0 to 6	--	--
	6 to 12	--	--
	12 to 24	--	--
	24 to 36	--	--
	36 to 48	--	--

APPENDIX D. DAILY ACTIVITY LOGS

Date, 09	No. of People	Area Tested	Status Start Time	Status Stop Time	Duration, min.	Operational Status	Operation Status - Comment	Track Method	Pattern	Field Conditions	
16 Nov	2	CALIBRATION LANES	1105	1340	155	INITIAL SETUP	INITIAL SETUP	GPS	LINEAR	SUNNY	MUDDY
	2	CALIBRATION LANES	1340	1510	90	COLLECTING DATA	COLLECT DATA	GPS	LINEAR	SUNNY	MUDDY
	2	BLIND TEST GRID	1510	1600	50	DAILY START, STOP	SET UP EQUIPMENT	GPS	LINEAR	SUNNY	MUDDY
	2	BLIND TEST GRID	1600	1640	40	DAILY START, STOP	EQUIPMENT BREAKDOWN	GPS	LINEAR	SUNNY	MUDDY
17 Nov	2	BLIND TEST GRID	750	815	25	DAILY START, STOP	SET UP EQUIPMENT	GPS	LINEAR	SUNNY	MUDDY
	2	BLIND TEST GRID	815	855	40	COLLECTING DATA	COLLECT DATA	GPS	LINEAR	SUNNY	MUDDY
	2	BLIND TEST GRID	855	930	35	DOWNTIME DUE TO EQUIP MAINT/CHECK	DATA CHECK	GPS	LINEAR	SUNNY	MUDDY
	2	OPEN FIELD	930	1145	135	DAILY START, STOP	SET UP EQUIPMENT	GPS	LINEAR	SUNNY	MUDDY
	2	OPEN FIELD	1145	1235	50	COLLECTING DATA	COLLECT DATA	GPS	LINEAR	SUNNY	MUDDY
	2	OPEN FIELD	1235	1330	55	BREAK/LUNCH	BREAK/LUNCH	GPS	LINEAR	SUNNY	MUDDY
	2	OPEN FIELD	1330	1520	110	COLLECTING DATA	COLLECT DATA	GPS	LINEAR	SUNNY	MUDDY
	2	CALIBRATION LANES	1520	1530	10	DAILY START, STOP	SET UP EQUIPMENT	GPS	LINEAR	SUNNY	MUDDY
	2	CALIBRATION LANES	1530	1550	20	COLLECTING DATA	COLLECT DATA	GPS	LINEAR	SUNNY	MUDDY
	2	CALIBRATION LANES	1550	1615	25	DAILY START, STOP	EQUIPMENT BREAKDOWN	GPS	LINEAR	SUNNY	MUDDY
18 Nov	2	OPEN FIELD	805	825	20	DAILY START, STOP	SET UP EQUIPMENT	GPS	LINEAR	SUNNY	MUDDY
	2	OPEN FIELD	825	1140	195	DOWNTIME DUE TO EQUIP MAINT/CHECK	CHARGING BATTERIES	GPS	LINEAR	SUNNY	MUDDY
	2	OPEN FIELD	1140	1330	110	COLLECTING DATA	COLLECT DATA	GPS	LINEAR	SUNNY	MUDDY
	2	OPEN FIELD	1330	1405	35	DOWNTIME DUE TO EQUIP MAINT/CHECK	DATA CHECK	GPS	LINEAR	SUNNY	MUDDY
	2	CALIBRATION LANES	1405	1440	35	COLLECTING DATA	COLLECT DATA	GPS	LINEAR	SUNNY	MUDDY
	2	CALIBRATION LANES	1440	1615	95	DEMOBILIZATION	DEMOBILIZATION	GPS	LINEAR	SUNNY	MUDDY

APPENDIX E. REFERENCES

1. Standardized UXO Technology Demonstration Site Handbook, DTC Project No. 8-CO-160-000-473, Report No. ATC-8349, March 2002.

2. Aberdeen Proving Ground Soil Survey Report, October 1998.

3. Data Summary, UXO Standardized Test Site: APG Soils Description, May 2002.

APPENDIX F. ABBREVIATIONS

ADST	=	Aberdeen Data Services Team
APG	=	Aberdeen Proving Ground
ATC	=	U.S. Army Aberdeen Test Center
ATSS	=	Aberdeen Test Support Services
BAR	=	background alarm rate
DMM	=	discarded military munitions
EMI	=	electromagnetic interference
EQT	=	Environmental Quality Technology
ERDC	=	U.S. Army Corps of Engineers Engineering Research and Development Center
EST	=	Eastern Standard Time
ESTCP	=	Environmental Security Technology Certification Program
GPS	=	Global Positioning System
GT	=	ground truth
HDSD	=	Homeland Defense and Sustainment Division
HEAT	=	high-explosive antitank
JPG	=	Jefferson Proving Ground
MM	=	military munitions
NS	=	nonstandard munition
POC	=	point of contact
QA	=	quality assurance
QC	=	quality control
ROC	=	receiver-operating characteristic
S	=	standard munition
SAIC	=	Science Applications International Corporation
SCEMP	=	Simplified Combined EMI Magnetometer Prototype
SERDP	=	Strategic Environmental Research and Development Program
USAEC	=	U.S. Army Environmental Command
UXO	=	unexploded ordnance
YPG	=	U.S. Army Yuma Proving Ground

(Page F-2 Blank)

APPENDIX G. DISTRIBUTION LIST

DTC Project No. 8-CO-160-UXO-021

Addressee	No. of Copies
Commander U.S. Army Aberdeen Test Center ATTN: TEDT-AT-SLE (Mr. J. Stephen McClung) 400 Colleran Road Aberdeen Proving Ground, MD 21005-5059	1
Program Manager SERDP/ESTCP Munitions Management ATTN: Mr. Herb Nelson 901 North Stuart Street, Suite 303 Arlington, VA 22203	1
Science Applications International Corporation (SAIC) ATTN: Mr. Robert Siegel 104 Clematis Avenue Waltham, MA 02453	1
Defense Technical Information Center 8725 John J. Kingman Road, Suite 0944 Fort Belvoir, VA 22060-6218	PDF

Secondary distribution is controlled by Program Manager, SERDP/ESTCP, Munitions Management, ATTN: Mr. Herb Nelson.